63-4-2

(5) 418 500

REPORT NO. RS-316

408 673

EFFECT OF ACCURACY LIMITATIONS IN THE AIRBORNE DIGITAL COMPUTER ON THE CONVERGENCE OF A STAGEWISE MIDCOURSE TRAJECTORY DETERMINATION PROCEDURE

13 June 1963

AEROSPACE GROUP

HUGHES

HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA

$4.60

(11) $4.60 (5) 418 500

(6) EFFECT OF ACCURACY LIMITATIONS IN THE AIRBORNE DIGITAL COMPUTER ON THE CONVERGENCE OF A STAGEWISE MIDCOURSE TRAJECTORY DETERMINATION PROCEDURE, (7) (9) NA

(10) by A. Holick,

(14) RS-316

(11) 13 June 1963, (12) 44p. (13) NA

(15) Contract AF04(695)210 (16) - (19) NA (20) U (21) NA

Hughes Aircraft Company · Culver City, California

ABSTRACT

The effect of round-off errors on the estimation of position and velocity in midcourse navigation ~~has been~~ WAS analyzed. The analysis is based on Kalman's approach to linear filtering and prediction. "Computation noise" appears to be an additional random force in the dynamic system, and may affect both convergence and equilibrium of the sequential estimation procedure significantly.

The analysis ~~has been~~ WAS applied to a satellite trajectory estimation system. Axes and area of the error ellipse (ellipse of concentration) ~~have been~~ WERE expressed in terms of word length, time interval between observations, and number of integration steps.

ACKNOWLEDGEMENT

The author gratefully acknowledges the encouragement given him on this paper by H. M. Breneman and the substantial suggestions by R. Deutch and N. A. Boehmer. C. J. Hasson performed the numerical computations.

The study has been performed under Air Force contract AF04(695)-210.

CONTENTS

ILLUSTRATIONS

TABLES

1. INTRODUCTION

Selection of the wordlength in digital computers used in guidance and navigation systems is infrequently considered to be a serious design problem.

The decision is most often based on whether 1 or 2 instructions should be stored in one memory word, and how many digits of the instruction word should be allotted to the order code and address. Index modification schemes for shortening the address may have more effect on the wordlength selection than accuracy requirements.

Reliability considerations increase the importance of wordlength studies significantly, especially if the digital computer is to be used on board a space vehicle and operates independently of ground communications. In optimizing the hardware for such a system, it is necessary to relate computer functions (i. e. accuracy) to the performance of the whole system, and to design the hardware so that intolerable degradations in performance resulting from failures are minimized or eliminated. A typical example of this realistic system optimization philosophy is the "Word Split Technique". It is a technique whereby a digital word is automatically split after a failure and the failing part is excluded from the operation[8,9]. Extensive studies at Hughes Aircraft Company[9,11] have shown that application of a "failure tolerance concept" instead of a pure redundancy concept may very likely lead to significant hardware savings. However, application of the failure tolerance concept to real time data requires a thorough analysis of the performance degradation caused by dropping the least significant word half.

The objective of this paper is to analyze the effect of computer accuracy limitations on system performance during midcourse navigation. The purpose is two fold:

a. To supply the tools for determining the error contributions of a (given) digital computer in midcourse navigation.

b. To establish whether degraded operational modes in extremely long range space missions may be tolerated .

Several papers have been published recently which describe the application of concepts from statistical filter theory to inflight determination of position and velocity of a (manned) space vehicle for the purpose of midcourse guidance. The spaceborne digital computer implements a dynamic time varying filter which weights the incoming observations in an "optimal" sense and produces an up to date optimal estimate of any desired set of state variables.

The basic theory quite frequently applied, is described in references (1) and (2) where it is shown that every solution of the variance equation converges to an equilibrium point, and that the equilibrium exists if certain conditions are satisfied. The variance equation is described as a stable computational method and is expected to be insensitive to round-off errors. Application of the theorems and hypothesis in (1) and (2) would always lead to the conclusion that computation errors are negligible compared to instrument anomalies. It seems questionable, however, to draw any conclusions about the propagation of round-off errors from a theory which does not permit such errors in the basic dynamic model.

In the following, we introduce "computation noise" in addition to the random forces and measurement uncertainties, and study its effect on the convergence and equilibrium of the estimation procedure.

2. SYSTEM MODEL

The fundamental relations for finding the best estimate of the message process in the linear dynamical system of the form

$$(1) \quad x(t+1) = \Phi(t+1,t)\, x(t) + \Gamma(t+1,t)\, w(t)$$

$$(2) \quad Z(t) = M(t)\, x(t) + v(t)$$

are given by

$$(3) \quad \hat{x}(t+1/t) = \Phi^{*}(t+1,t)\, \hat{x}(t/t-1) + \Delta^{*}(t)\, Z(t)$$

$$(4) \quad \Phi^{*}(t+1,t) = \Phi(t+1,t) - \Delta^{*}(t)\, M(t)$$

$$(5) \quad \Delta^{*}(t) = \Phi(t+1,t)\, P(t/t-1)\, M^{T}(t) \left[M(t)\, P(t/t-1) M^{T}(t) + R(t) \right]^{-1}$$

$$(6) \quad P(t+1,t) = \Phi(t+1,t) \left\{ P(t/t-1) - \left[P(t/t-1)\, M^{T}(t) \right] \left[M(t)\, P(t/t-1)\, M(t) + R(t) \right]^{-1} \left[M(t)\, P(t/t-1) \right] \right\} \Phi^{T}(t+1,t) + \Gamma(t+1,t)\, Q(t)\, \Gamma^{T}(t+1,t)$$

where

$\Phi(t+1,t)$	=	state transition matrix
$x(t)$	=	state vector
$Z(t)$	=	vector of measurements
$M(t)$	=	transformation relating the observables to the state vector
$\Gamma(t+1,t)$	=	transition matrix for the random force $w(t)$
$v(t)$	=	measurement uncertainty
$w(t)$	=	random force
$\hat{x}(t/t-1)$	=	estimated state vector based on past observation

$\Delta^*(t)$ = optimum filter

$P(t+1, t)$ = covariance matrix of the estimation error

$R(t)$ = covariance matrix of the measurement uncertainties $v(t)$

$Q(t)$ = covariance matrix of the random force $w(t)$

It is assumed that

$$(7) \quad E\left[w(t)\right] = E\left[v(t)\right] = 0$$

$$(8) \quad E\left[w(t)\, w^T(t)\right] = Q(t)$$

$$(9) \quad E\left[v(t)\, v^T(t)\right] = R(t)$$

For the derivation of the formulas see reference (1) and (2).

3. DEFINITION OF COMPUTATION NOISE

No digital computing procedure or device can perform the operations which are its "elementary" operations rigorously and faultlessly because of the finite length of a "digital word". Each time an elementary operation is performed, a perturbation is introduced which will cause a parameter to deviate from its ideal value. The non ideal elementary operations will be called "pseudo operations". They form a constantly renewed source of contamination, and their influence increases with the number of elementary operations that have to be performed. They are therefore especially important in long computations, involving many such operations. Long computations will undoubtedly be normal for midcourse estimation and decision making procedures. The decisive factor that controls their effect is some kind of stability phenomenon. And it is the stability of the approximant procedure and not of the strict procedure that matters. Estimation procedures which are theoretically identical (asymptotically), but differ in the number and sequence of pseudo operations will be characterized by a different speed of convergence to a different equilibrium point.

Notation for pseudo operations:

$$\overline{A} = A + \eta \quad \text{where} \quad |\eta| \leq 2^{-(K+1)} \quad (K = \text{wordlength})$$

$$\overline{A} \oplus \overline{B} = \overline{A} + \overline{B} + \eta_a \qquad |\eta_a| \leq 2^{-(K+1)}$$

$$\overline{A} \times \overline{B} = \overline{A} \,.\, \overline{B} + \eta_m \qquad |\eta_m| \leq 2^{-(K+1)}$$

$$\overline{A} \div \overline{B} = \overline{A}/\overline{B} + \eta_d \qquad |\eta_d| \leq 2^{-(K+1)}$$

The η's are random variables and assumed to be uniformly distributed between $\pm\, 2^{-(K+1)}$. In other words, all bit configurations are assumed to be equally likely for $j \geq k$, where j designates the truncated digits.

It is remarked that scaling is also a pseudo operation. It is identical to a multiplication or division, but is carried out with a multiplier or divisor outside of the range of digital numbers in a fixed point organization.

The pseudo operations considered so far are the "noisy operations" in the failure free mode. These can be characterized by a uniform error distribution which converges rapidly to the normal distribution as the number of pseudo operations performed in sequence increases. But if permanent hardware failures are "allowed to occur" in long range space missions, the computer may very well degrade its performance[8,9,10,11]. The statistics of pseudo operations in degraded modes (time of occurrence is a random variable) is not analyzed here.

4. EFFECT OF COMPUTATION NOISE ON THE ESTIMATION ERROR

The estimate of the state vector is computed from

$$\hat{x}(t+1/t) = \Phi_D^*(t+1,t) \times \hat{x}(t/t-1) \oplus \Delta_D^*(t) \times Z(t)$$

where the index D indicates that the numerical values of the matrices Φ^* and Δ^* are erroneous:

$$(10) \quad \Phi_D^* = \Phi^* + \delta\Phi^*$$

$$(11) \quad \Delta_D^* = \Delta^* + \delta\Delta^*$$

The estimation error in step (t+1,t) caused only by computation noise is therefore in a first approximation:

$$(12) \quad \delta\hat{x}(t+1,t) = \left[\bar{\Phi} \times \bar{\hat{x}} - \Phi \,.\, \hat{x}\right] + \left[\bar{\Delta}^* \times \bar{Z} - \Delta^* \,.\, Z\right] - \left[\bar{\Delta}^* \times \bar{M} \times \bar{\hat{x}} - \Delta^* \,.\, M \,.\, \hat{x}\right] + \delta\Phi\hat{x} + \delta\Delta^* Z - \delta\Delta^* M\hat{x}$$

The total estimation error in step (t+1,t) is:

$$(13) \quad \tilde{x}_D(t+1/t) = \tilde{x}(t+1/t) + \delta\tilde{x}(t+1,t)$$

For simplification it is assumed in the following that the errors introduced by pseudo operations in calculating $\hat{x}(t+1,t)$ are negligible against those introduced in computing Φ^* and Δ^*. Thus,

$$(14) \quad \delta\hat{x}(t+1,t) \cong \delta\Phi\hat{x} + \delta\Delta^* Z - \delta\Delta^* M\hat{x}$$

Some properties of $\delta\Phi$ and $\delta\Delta^*$ which follow from the algebra of pseudo-operations are listed below:

$$
\begin{aligned}
(15)\quad E\left[\delta\Phi\right] &= \left[\delta\Delta^*\right] = 0 \\
E\left[\delta\Phi\, x\right] &= E\left[\delta\Phi\right] E\left[\tilde{x}\right] = 0 \\
E\left[\delta\Delta^* \tilde{x}\right] &= 0 \\
E\left[\delta\Phi^j \tilde{x}^K\right] &= E\left[\delta\Phi^j\right] E\left[\tilde{x}^K\right] \\
E\left[\delta\Delta^{*j} \tilde{x}^K\right] &= E\left[\delta\Delta^{*j}\right] E\left[\tilde{x}^K\right] \\
E\left[\delta\Phi\,\delta\Delta^*\right] &\neq E\left[\delta\Phi\right] E\left[\delta\Delta^*\right] \\
E\left[\delta\Phi^i_{t+1}\,\delta\Phi^K_t\right] &= E\left[\delta\Phi^i_{t+1}\right] E\left[\delta\Phi^K_t\right], \text{ similar for } \delta\Delta^*
\end{aligned}
$$

Introducing (2) and (13) into (12) we obtain

$$\delta\hat{x}(t+1,t) = \delta\Phi\,\hat{x}(t/t-1) + \delta\Delta^* M\tilde{x}_D(t/t-1)$$

where

$$\hat{x}(t/t-1) = x(t) + \tilde{x}(t/t-1) + \delta\hat{x}(t,t-1)$$

and

$$\tilde{x}_D(t/t-1) = \tilde{x}(t/t-1) + \delta\hat{x}(t,t-1)$$

or

$$\tilde{x}_D(t/t-1) = \delta\Delta^* M\tilde{x}_D(t-1/t-2) + \tilde{x}(t/t-1) + \delta\Phi\hat{x}(t-1/t-2)$$

Repeated application of the recursion formulas above leads to:

$$(16)\quad \delta\hat{x}(t+1,t) = \sum_i (\delta\Phi + \delta\Delta^* M)^{i-1} \cdot \delta\Phi\, x(t-i) + \sum_i (\delta\Phi + \delta\Delta^* M)^{i}\, \tilde{x}(t-i/t-i-1) + (\delta\Phi + \delta\Delta^* M)^{n}\, \delta\hat{x}(t-n)$$

The effect of errors made in previous computation (observation) intervals is gradually forgotten as $t \to \infty$; but the speed of convergence depends on the behavior of the state vector (first term in 16) and estimation error (second term in 16) as a function of time. It depends also, of course, on the magnitude of $\delta\Phi$ and $\delta\Delta^* M$. Assuming both to be moderately small, the last term in (16) can be neglected.

Because of the dependency of $\delta\hat{x}$ on the state vector, convergence and equilibrium are not guaranteed anymore by the conditions given in [1] and [2] as it will be shown in a simple example in section 5.

It may be sufficient to consider only first order terms:

$$(17)\quad \delta\hat{x}(t+1,t) \cong \delta\Phi\, x(t) + (\delta\Phi + \delta\Delta^* M)\, \tilde{x}(t/t-1)$$

5. COVARIANCE MATRIX OF ESTIMATION ERROR (VARIANCE EQUATION

Applying the expectation operator in a straight forward manner according to

$$P(t+1) = E\left[\tilde{x}_D(t+1/t)\,\tilde{x}_D^T(t+1/t)\right]$$

and considering the relations given in (13) and (16) results in

$$(18)\quad P_D(t+1) = \Phi_D(t+1,t)\left[P_I + \delta P_I\right]\Phi_D^T(t+1,t) + \Gamma(t+1,t)\,Q(t)\,\Gamma^T(t+1,t) + \\ + E\left[\delta\Phi\,\hat{x}(t/t-1)\,\hat{x}^T(t/t-1)\,\delta\Phi^T\right] + \\ + E\left[\delta\Delta^*\,M\tilde{x}(t/t-1)\,\tilde{x}^T(t/t-1)\,M^T\delta\Delta^{*T}\right] + \\ + E\left[\delta\Delta^*\,M\delta\hat{x}(t)\,\delta\hat{x}^T(t)\,M^T\delta\Delta^{*T}\right]$$

where

$$P_I = P(t) - P(t)\,M^T\left[MPM^T - R(t)\right]^{-1} MP(t)$$

Round-off errors contribute to $P_D(t+1)$ in two ways:

1. In computing the estimator $\hat{x}(t+1)$ according to equation (3) we get the terms

$$E\left[\delta\Phi\hat{x}\,\hat{x}^T\cdot\delta\Phi^T\right]$$

$$E\left[\delta\Delta^*\,M\tilde{x}\,\tilde{x}^T\,M^T\delta\Delta^{*T}\right]$$

$$E\left[\delta\Delta^*\,M\delta\hat{x}\,\delta\hat{x}^T\,M^T\delta\Delta^{*T}\right]$$

2. In computing the covariance matrix P(t+1) according to the Variance Equation (6) we get the terms

$$\left[\delta\Phi P_I \delta\Phi^T \right] \quad \text{and} \quad \left[\Phi \delta P_I \Phi^T \right]$$

The covariance matrix of the estimation error is not a deterministic function of the time anymore. Its elements are random variables because of the randomness of $\delta\Phi$ and δP_I. Taking the expectation (bias term), we obtain

$$(19)\quad E\left[P_D(t+1)\right] = P(t+1) + E\left[\delta\Phi P_I \delta\Phi^T\right] + E\left[\delta\Phi \hat{x}(t/t-1)\, \hat{x}^T(t/t-1)\, \delta\Phi^T\right] + E\left[\delta\Delta^* M\tilde{x}(t)\, \tilde{x}^T(t)\, M^T \delta\Delta^{*T}\right] + E\left[\delta\Delta^* M\delta\hat{x}(t,t-1)\, \delta\hat{x}^T(t,t-1)\, M^T \delta\Delta^{*T}\right]$$

where

$$P_I = P(t) - P(t)\, M^T \left\{ MP(t)\, M^T + R(t) \right\}^{-1} MP(t)$$

It will normally be sufficient to use (20) instead of (19):

$$(20)\quad E\left[P_D(t+1)\right] \cong P(t+1) + E\left[\delta\Phi\, \hat{x}(t/t-1)\, \hat{x}^T(t/t-1)\, \delta\Phi^T\right]$$

The second term on the right is a diagonal matrix with the elements:

$$(21)\quad \delta P_{ii}(t+1) \cong \sum_j E\left[\delta\Phi_{ij}^2\, \hat{x}_j^2\right]$$

$$\delta P_{ij}(t+1) = 0 \quad \text{for } i \neq j$$

In a first approximation it will suffice to apply

$$E\left[\delta\Phi_{ij}^{2}\,\hat{x}_j^2\right] \cong x_j^2 \, . \, E\left[\delta\Phi_{ij}^{2}\right].$$

Thus,

$$(22) \quad \delta P_{ii}(t+1) \cong \sum_j x_j^2 \, E\left[\delta\Phi_{ij}^{2}\right]$$

<u>One Dimensional Example</u>

The system is determined by

$$x(t+1) = \zeta(t+1,t)\, x(t) + u(t)$$

$$Z(t) = x(t) + v(t)$$

where

$$Eu^2 = q \text{ and } E\,v^2 = r$$

The expectation of the variance at time t+1 is according to equation (19) given by

$$E\left[p(t+1)\right. = E\left[\sigma^2(t+1)\right. = \frac{p(t)r}{p(t)+r}\;\zeta^2(t+1,t) + \left[\frac{p(t)r}{p(t)+r} + x^2(t)\right]$$

$$E\left[\delta\zeta^2\right] + q$$

The equilibrium point can be obtained by setting $p(t+1) = p(t) = \overline{p}$:

$$\overline{p} = \frac{\overline{p}r}{p+r}\;(\zeta^2 + E\delta\zeta^2) + E\left[\hat{x}^2(t)\,\delta\zeta^2\right]$$

or (for constant ζ):

$$\overline{p} = \frac{1}{2}\,S \pm \sqrt{\frac{S^2}{4} + r\,(q + x^2(t)\;E\delta\zeta^2)}$$

where

$$S = r(\zeta^2 + E\delta\zeta^2 - 1) + q + x^2(t)\, E\delta\zeta^2$$

Existence and numerical value of the equilibrium depends obviously on the state variable $x(t)$ and, therefore, on ζ.

For $\zeta > 1$:

$$\lim_{t \to \infty} \overline{p} = \infty$$

For $\zeta = 1$:

$\lim_{t \to \infty} \overline{p}$ exists, but can be intolerably large if the coordinate system cannot be chosen properly.

For $\zeta < 1$:

$\lim_{t \to \infty} \overline{p}$ does not depend on the state variable $x(t)$.

If ζ is a function of time, then the conditions for existence of equilibrium are surprisingly reduced: ζ can be larger than 1, if, for instance, $\zeta(t)$ is periodic and the time average is ≤ 1. But it is not sufficient to require only that, as stated in reference 2

$$0 < \delta \leq /\zeta(t+1, t)/ \leq \beta, < \infty$$

6. MODIFIED SYSTEM MODEL

Inspection of Equation (3) shows that propagation of injection errors and unperturbed motion of the vehicle is somewhat intermixed. The state transition matrix determines the propagation of injection errors as well as the unperturbed motion of the vehicle.

If we define two transition matrices, one for the perturbations and one for the reference trajectory, then the estimation problem could be separated from the problem of integrating the equations of motion. A system model is shown in Figure 1:

1. ${}_{E}\Phi(t_1, t)$ is the transition matrix for the perturbations and determines the effect of injection errors at an arbitrary time t_1. The reference trajectory can be assumed to be recomputed every time our knowledge of the injection conditions improves.
2. ${}_{M}\Phi(t_1, t)$ is the transition matrix for the state vector X (t) and determines the unperturbed state of the system at an arbitrary time t_1. The elements of ${}_{M}\Phi$ may be changed every time our knowledge of the actual trajectory improves (the perturbations tend to zero if the estimation procedure converges).

We estimate now the perturbations x(t) and perform the computation of the state vector X (t) outside of the estimation loop. We assume that knowledge of X (t) is required throughout the mission to decide if corrective maneuvers should be made, and for a more accurate (not linearized) representation of the system model.

The modified error equations are readily found with

$$ {}_{M}\Phi_{D} = {}_{M}\Phi + \delta\, {}_{M}\Phi $$

$$ {}_{E}\Phi_{D} = {}_{E}\Phi + \delta\, {}_{E}\Phi $$

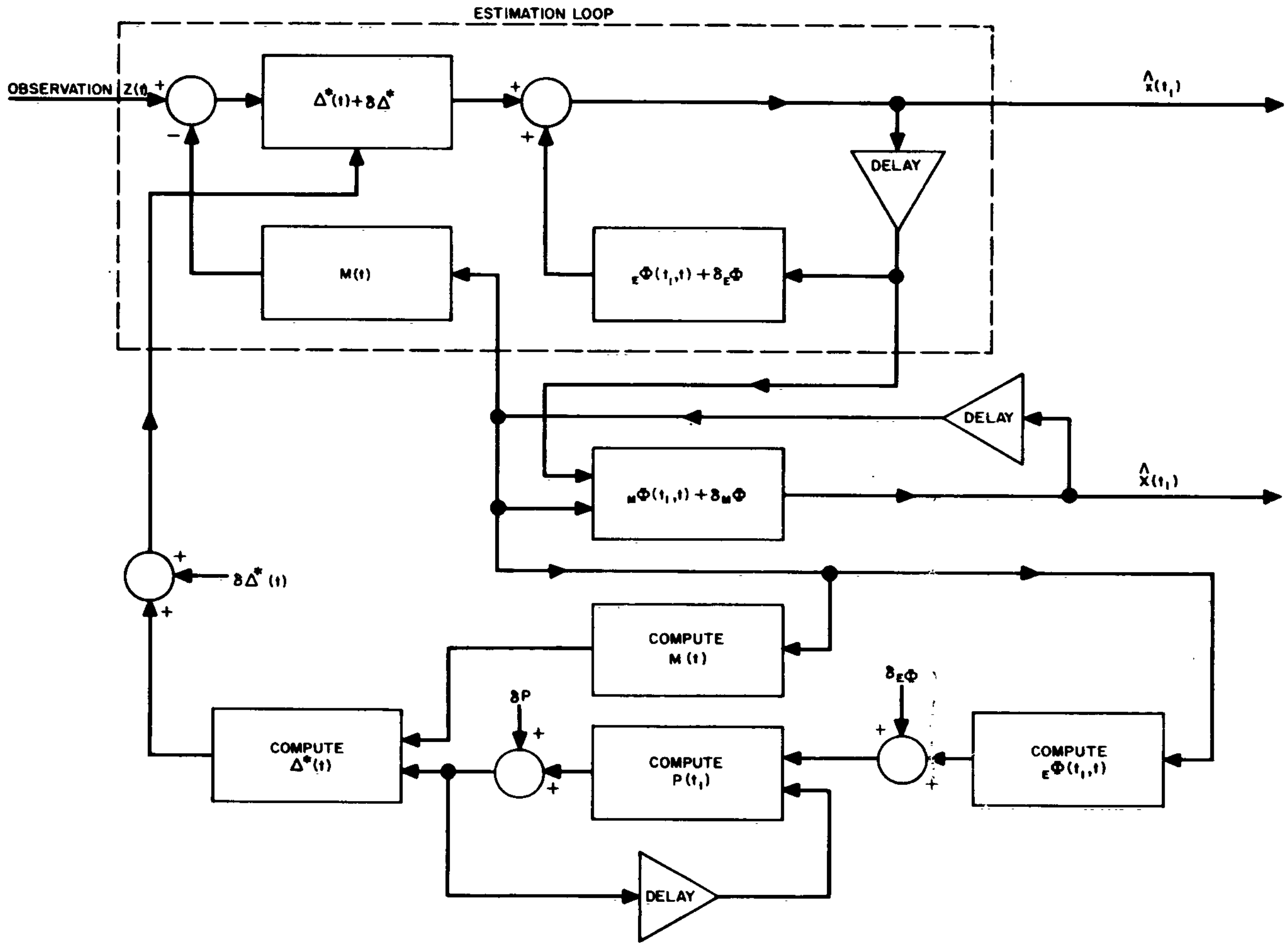

Figure 1. A trajectory estimation system with computation noise; estimation of the perturbation vector $\hat{x}(t_1)$.

We note that the round-off errors which are generated during numerical integration of the equations of motion still contribute to the estimation error $\tilde{x}_D(t)$:

$$(23) \qquad \tilde{x}_D(t+1) = \tilde{x}(t+1/t) + \delta\hat{x}(t+1)$$

where

$$\delta\hat{x}(t+1) = \delta_E\Phi\hat{x}(t/t-1) + \delta\Delta^* M\tilde{x}_D(t/t-1) + \delta_M\Phi\hat{X}(t)$$

or

$$\delta\hat{x}(t+1) = (\delta_E\Phi + \delta_M\Phi)\hat{x}(t) + \delta_M\Phi X(t) + \delta\Delta^* M\tilde{x}_D(t/t-1)$$

Repeated application of the recursion formulas above leads to:

$$(24) \qquad \delta\hat{x}_D(t) = \sum_i (\delta_E\Phi + \delta_M\Phi + \delta\Delta^* M)^{i-1}(\delta_M\Phi + \delta_E\Phi)x(t-i) +$$

$$+ \sum_i (\delta_E\Phi + \delta_M\Phi + \delta\Delta^* M)^i \tilde{x}(t-i) +$$

$$+ \sum_i (\delta_E\Phi + \delta_M\Phi + \delta\Delta^* M)^{i-1} \delta_M\Phi X(t-i)$$

and shows that the propagation of round-off errors from previous observation intervals can be neglected, if $\delta_M\Phi$, $\delta_E\Phi$ and $\delta\Delta$ are sufficiently small. In a first approximation,

$$(25) \qquad \delta\hat{x}_D(t) \cong (\delta_M\Phi + \delta_E\Phi)x(t-1) + (\delta_M\Phi + \delta_E\Phi + \delta\Delta^* M)\tilde{x}(t-1) +$$

$$+ \delta_M\Phi X(t-1)$$

The modified covariance matrix is given by

$$E\left[P_D(t+1)\right] = E\left[E\left[\tilde{x}_D(t+1/t)\,\tilde{x}_D(t+1/t)\right]\right] =$$

$$= P(t+1) + E\left[\delta_E\Phi P_I(t)\,\delta_E\Phi^T\right] + E\left[\delta_M\Phi X(t)\,X^T(t)\,\delta_M\Phi^T\right] +$$

$$+ E\left[(\delta_M\Phi + \delta_E\Phi)\,\hat{x}(t)\,\hat{x}^T(t)\,(\delta_M\Phi^T + \delta_E\Phi^T)\right] + \tag{26}$$

$$+ E\left[\delta\Delta^* M\,\tilde{x}_D(t)\,\tilde{x}_D^T(t)\,M^T\,\delta\Delta^{*T}\right]$$

Round-off errors in the filter operation $\Delta^*(t)$ can safely be disregarded if the regular estimation error $\tilde{x}(t)$ converges in a mean square sense, or, if

$$\lim_{N\to\infty} E\left[\tilde{x}\,\tilde{x}^T\right] = 0.$$

We write therefore

$$E\left[P_D(t+1)\right] \cong P(t+1) + E\left[\delta_M\Phi X(t)\,X^T(t)\,\delta_M\Phi^T\right] +$$

$$+ E\left[(\delta_M\Phi + \delta_E\Phi)\,\hat{x}(t)\,\hat{x}^T(t)\,(\delta_M\Phi^T + \delta_E\Phi^T)\right] \tag{27}$$

With the same arguments as in Section 5, we obtain for the error in element $P_{D,\,ii}$

$$\boxed{\delta P_{ii}(t+1) \cong \sum_j x_j^2(t)\,E\left[\delta_M\Phi_{ij}^2 + \delta_E\Phi_{ij}^2\right] + \sum_j X_j^2(t)\,E\left[\delta_M\Phi_{ij}^2\right] \qquad (28)}$$

We are now in the position to discuss the following case: Suppose one would like to reduce the effect of round-off errors by estimating the time independent injection errors $x(t_o)$ rather than the current perturbations $x(t)$. The state transition matrix is here

$${}_E\Phi(t_1, t_o) = 1 \quad \text{and} \quad \delta_E\Phi = 0$$

But in computing the present state X(t) each time a new measurement is made or each time a decision for a corrective maneuver is required, the equations of motion have to be integrated from t_o to t. The variance of $\delta_M\Phi(t, t_o)$ increases significantly because of the increased integration time (Section 7) and may finally become dominant. This method seems favourable only if $_M\Phi$ does not have to be computed with numerical integration.

If we are forced by some reason to determine $X(t_1)$ by integrating the equations of motion, then it seems advisable to modify our estimation procedure so that $({}_M\Phi + \delta_M\Phi)$ becomes part of the estimation loop (Figure 2).

The covariance matrix of the estimation error is given here by the diagonal matrix:

$$E\left[P_D(t+1)\right] \cong E\left[\delta_M\Phi\ X(t)\ X^T(t)\ \delta_M\Phi\right] = E\left[\delta X(t_1)\ \delta X^T(t_1)\right] \tag{29}$$

where

$$E\left[\delta X(t_1)\ \delta X^T(t_1)\right] = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

The σ_i^2 are equal to Var $\left[u_i(t_M)\right]$ and given in Table 2.

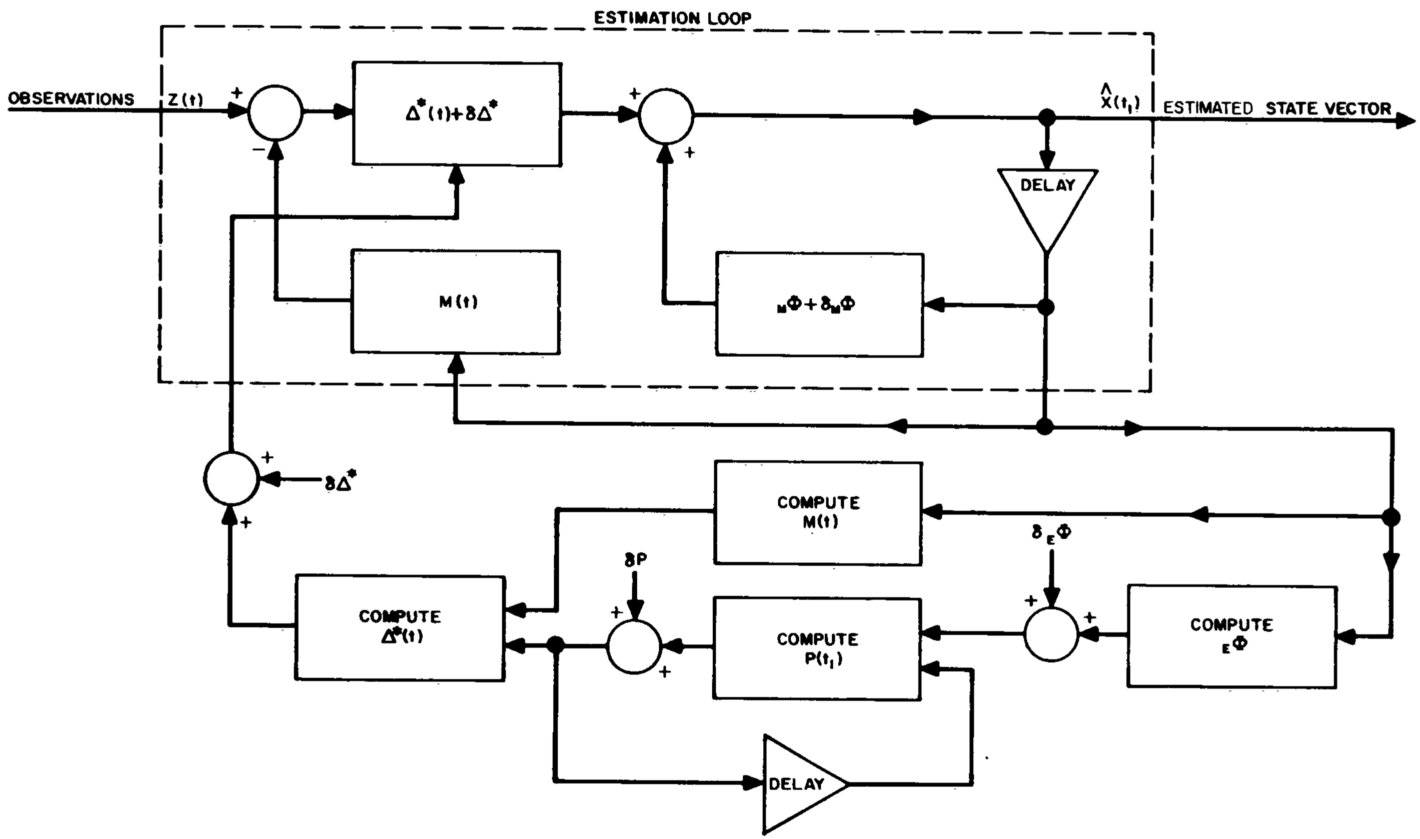

Figure 2. A trajectory estimation system with computation noise; estimation of the system state vector $\hat{X}(t_1)$.

7. EFFECT OF COMPUTATION NOISE ON THE STATE TRANSITION MATRIX

7.1 ERRONEOUS PREDICTION OF THE PERTURBATIONS

A method for computing the elements of the state transition matrix, $_E\Phi$, has been described in (5). The elements are found there with six-fold numerical integration of the perturbation equations

$$\dot{x}(t) = F(t)\, x(t)$$

under the appropriate initial conditions. We choose this method as an example for studying the round-off error propagation, because it is quite generally applicable and requires especially high computation accuracy (worst case example).

We are not interested here in a complete round-off error analysis where

the method of integration used,

the form of the equation to be integrated, and

the program organization

would have to be considered in detail. We restrict our study to a symmetric round-off, Heun's method of integration, and assume optimal scaling. We also do not analyze the effect of the equation form on error propagation. However, we keep in mind that canonical transformations (separation of the variables) could probably always be used to reduce the accumulated round-off error. Each equation of the system could be integrated separately by means of the trapezoidal rule and the accumulated round-off error could be made proportional to $\frac{1}{\Delta t}$ or $\frac{t_M}{N}$ for each variable.

We intend to derive relations between the statistical moments of the round-off errors in the state transition matrix and computation parameters like wordlength and number of integration steps.

It seems desirable for our purpose to obtain mean and variance in closed form rather than by simulation methods. A circular orbit with the state variables r, θ , $\dot{r}$, $\dot{\theta}$ was chosen, therefore, as an example. The error Equations (33) and (34) can be solved in closed form.

Generalization to non circular orbits and 6 state variables can readily be accomplished (Appendix A) but not in closed form. We would have to find polynomial approximations for the solution of the adjoint system (34). However, the results in this section can also be used to approximate the round-off error in the general case.

The error variables are defined as the deviations of the actual from the reference trajectory:

$$x_1 = r - r_R \quad , \quad x_2 = \theta - \theta_R \quad , \quad x_3 = \dot{r} - \dot{r}_R \quad , \quad x_4 = \dot{\theta} - \dot{\theta}_R \quad ,$$

where r and θ are the polar coordinates of the vehicle with the origin in the center of attraction.

From the equations of motion

$$\ddot{r} = - \frac{K}{r^2} + r\,\dot{\theta}^2 \qquad\qquad (30)$$

$$\ddot{\theta} = - \dot{\theta}\,\frac{\dot{r}}{r}$$

we find readily the perturbation equations (perturbations are assumed to be caused by injection errors only) with Taylor expansion of the right side of (30). The perturbation equations are

$$\dot{x}_1(t) = x_3(t)$$

$$\dot{x}_2(t) = x_4(t)$$

$$\dot{x}_3(t) = 3\omega^2\,x_1(t) + 2\,r\,\omega\,x_4(t) \quad ; \qquad \omega = \dot{\theta} = \text{constant} \qquad (31)$$

$$\dot{x}_4(t) = - \frac{\omega}{r}\,x_3(t)$$

and could be integrated in closed form. But we integrate numerically with Heun's method* [4] to study the propagation of round-off errors. We compute first the auxiliary values

*Heun's method is less accurate than, for instance, the Runge Kutta method, but simpler to analyze. Its accuracy seems sufficient for our objective "to study the effect of wordlength on the accuracy of the state transition matrix."

$$x_i^*(t_j) = x_i(t_{j-1}) + \Delta t\, f_i\left[x_1(t_{j-1}), \ldots, x_n(t_{j-1})\right]$$

and then

$$x_i(t_j) = x_i(t_{j-1}) + \frac{\Delta t}{2} f_i\left\{\left[x_1(t_{j-1}), \ldots, x_n(t_{j-1})\right] + f_i\left[x_1^*(t_j), \ldots, x_n^*(t_j)\right]\right\}$$

$$i = 1, \ldots, n \text{ and } j = 1, \ldots, N$$

The determination of round-off errors can be considerably simplified, if we assume that Δt is a small quantity so that multiplication by Δt shifts the doubtful figures in $f_i(x_1 \ldots x_n)$ into those which are dropped in the last round-off step. We have the identity

$$\overline{\Delta t \cdot f\left[\overline{x}_1(t_{j-1}) \ldots \overline{x}_n(t_{j-1})\right]} = \overline{\Delta t \cdot \overline{f\left[\overline{x}_1(t_{j-1}) \ldots \overline{x}_n(t_{j-1})\right]}}$$

where the bars indicate round-off values.

If we are not interested in higher powers of Δt, then we may write for the j^{th} step:

$$\overline{x}_{1j} = \overline{x}_{1\,j-1} + \Delta t\, \overline{x}_{3\,j-1} + 2\,\epsilon_1$$

$$\overline{x}_{2j} = \overline{x}_{2\,j-1} + \Delta t\, \overline{x}_{4\,j-1} + 2\,\epsilon_2$$

$$\overline{x}_{3j} = \overline{x}_{3\,j-1} + \Delta t\left(3\omega^2\, \overline{x}_{1\,j-1} + 2\,r\omega\, \overline{x}_{4\,j-1}\right) + 2\,\epsilon_3 \tag{32}$$

$$\overline{x}_{4j} = \overline{x}_{4\,j-1} - \Delta t\, \frac{\omega}{r}\, \overline{x}_{3\,j-1} + 2\,\epsilon_4$$

which we have to compare with

$$x_{1j} = x_{1\,j-1} + \Delta t\, x_{3\,j-1}$$

- - - - - - - - - - - - - -

$$x_{4j} = x_{4\,j-1} - \Delta t\, \frac{\omega}{r}\, x_{3\,j-1}$$

Introducing the error variables

$$u_{ij} = x_{ij} - \overline{x}_{ij}$$

we can write the error equation for step j

$$u_{1j} = u_{1j-1} + \Delta t \left[x_{3\,j-1} - \overline{x}_{3\,j-1} \right] - 2\,\epsilon_1$$

- -

$$u_{4\,j} = u_{4\,j-1} + \Delta t \left[- (\frac{\omega}{r})\, x_{3\,j-1} + (\frac{\omega}{r})\, \overline{x}_{3\,j-1} \right] - 2\,\epsilon_4$$

or

$$\begin{aligned}
u_{1\,j} &= u_{1\,j-1} + \Delta t\, u_{3\,j-1} - 2\,\epsilon_1 \\
u_{2\,j} &= u_{2\,j-1} + \Delta t\, u_{4\,j-1} - 2\,\epsilon_2 \\
u_{3\,j} &= u_{3\,j-1} + \Delta t \left(3\,\omega^2\, u_{1\,j-1} + 2\,r\,\omega\, u_{4\,j-1} \right) - 2\,\epsilon_3 \\
u_{4\,j} &= u_{4\,j-1} - \Delta t \left(\frac{\omega}{r}\, u_{3\,j-1} \right) - 2\,\epsilon_4
\end{aligned} \tag{33}$$

To solve Equation System (33) we introduce the adjoint system

$$\begin{aligned}
\lambda_{1j} &= \lambda_{1\,j-1} - \Delta t \cdot 3\,\omega^2\, \lambda_{3\,j} \\
\lambda_{2\,j} &= \lambda_{2\,j-1} \\
\lambda_{3\,j} &= \lambda_{3\,j-1} - \Delta t \left(\lambda_{1\,j} - \frac{\omega}{r}\, \lambda_{4\,j} \right) \\
\lambda_{4\,j} &= \lambda_{4\,j-1} - \Delta t \left(\lambda_{2\,j} + 2\,r\,\omega\, \lambda_{3\,j} \right)
\end{aligned} \tag{34}$$

and infer

$$(\lambda_{1j}u_{1j} - \lambda_{1j-1}u_{1j-1}) + (\lambda_{2j}u_{2j} - \lambda_{2j-1}u_{2j-1}) + (\lambda_{3j}u_{3j} - \lambda_{3j-1}u_{3j-1}) +$$

$$+ (\lambda_{4j}u_{4j} - \lambda_{4j-1}u_{4j-1}) = -2(\lambda_{1j}\epsilon_1 + \lambda_{2j}\epsilon_2 + \lambda_{3j}\epsilon_3 + \lambda_{4j}\epsilon_4)$$

Summing over j from 1 to N we obtain

$$\lambda_1(t_M)\,u_1(t_M) + \lambda_2(t_M)\,u_2(t_M) + \lambda_3(t_M)\,u_3(t_M) + \lambda_4(t_M)\,u_4(t_M)$$

$$= -2\sum_{j=1}^{N}(\lambda_{1j}\epsilon_1 + \lambda_{2j}\epsilon_2 + \lambda_{3j}\epsilon_3 + \lambda_{4j}\epsilon_4) \tag{35}$$

where

$t_M = N\Delta t$ = total time of integration

Δt = integration step

Treating the ϵ's as independently and uniformly distributed, we can write down the variance of (35)

$$\mathrm{Var}\left[\sum_{i=1}^{4}\lambda_i(t_M)\,u_i(t_M)\right] = 4\sum_{j=1}^{N}\left(\lambda_{1j}^2 + \lambda_{2j}^2 + \lambda_{3j}^2 + \lambda_{4j}^2\right)\sigma^2 \tag{36}$$

where σ^2 designates the round-off error variance in one integration step. The sums on the right can be looked upon as Rieman sums and can be replaced by the integrals

$$\mathrm{Var}\left[\sum_{i=1}^{4}\lambda_i(t_M)\,u_i(t_M)\right] = \frac{4\sigma^2}{\Delta t}\int_0^{t_M}\left(\lambda_1^2(t) + \lambda_2^2(t) + \lambda_3^2(t) + \lambda_4^2(t)\right)\,dt \tag{37}$$

In order to get the variances in u_i separately, one has to impose on $\lambda_i(t)$ the terminal conditions:

$$
\begin{array}{ll}
1. & \lambda_1(t_M) = 1 \quad , \quad \lambda_2(t_M) = \lambda_3(t_M) = \lambda_4(t_M) = 0 \\
2. & \lambda_2(t_M) = 1 \quad , \quad \lambda_1(t_M) = \lambda_3(t_M) = \lambda_4(t_M) = 0 \\
3. & \lambda_3(t_M) = 1 \quad , \quad \lambda_1(t_M) = \lambda_2(t_M) = \lambda_4(t_M) = 0 \\
4. & \lambda_4(t_M) = 1 \quad , \quad \lambda_1(t_M) = \lambda_2(t_M) = \lambda_3(t_M) = 0
\end{array}
\tag{38}
$$

SOLUTION OF THE ADJOINT SYSTEM

We can write the adjoint system in differential form:

$$
\begin{aligned}
\dot{\lambda}_1 &= -3\,\omega^2 \lambda_3 \\
\dot{\lambda}_2 &= 0 \\
\dot{\lambda}_3 &= -\lambda_1 + \frac{\omega}{r} \lambda_4 \\
\dot{\lambda}_4 &= -\lambda_2 - 2\,r\,\omega\,\lambda_3
\end{aligned}
\tag{39}
$$

and obtain

$$
\begin{aligned}
\lambda_1 &= \frac{\omega}{r} C_4 - \frac{3\omega}{r} C_1 t - 3\omega C_2 e^{\omega t} + 3\omega C_3 e^{-\omega t} \\
\lambda_2 &= C_1 \\
\lambda_3 &= \frac{C_1}{r\omega} + C_2 e^{\omega t} + C_3 e^{-\omega t} \\
\lambda_4 &= C_4 - 3 C_1 t - 2 r C_2 e^{\omega t} + 2 r C_3 e^{-\omega t}
\end{aligned}
\tag{40}
$$

For the four sets of terminal conditions we get four sets of solutions (Table 1).

Terminal Conditions				Solution of the Adjoint System
$\lambda_1(t_M)$	$\lambda_2(t_M)$	$\lambda_3(t_M)$	$\lambda_4(t_M)$	
1	0	0	0	$\lambda_1(t) = 3 \cos h\omega(t_M - t) - 2$ $\lambda_2(t) = 0$ $\lambda_3(t) = \frac{1}{\omega} \sin h\,\omega(t_M - t)$ $\lambda_4(t) = \frac{2r}{\omega} \left[\cos h\,\omega(t_M - t) - 1\right]$
0	1	0	0	$\lambda_1(t) = \frac{3\omega}{r}(t_M - t) - \frac{3}{r} \sin h\omega(t_M - t)$ $\lambda_2(t) = 1$ $\lambda_3(t) = \frac{1}{r\omega} \left[1 - \cos h\omega(t_M - t)\right]$ $\lambda_4(t) = 3(t_M - t) - \frac{2}{\omega} \sin h\,\omega(t_M - t)$
0	0	1	0	$\lambda_1(t) = 3\,\omega \sin h\omega(t_M - t)$ $\lambda_2(t) = 0$ $\lambda_3(t) = \cos h\omega(t_M - t)$ $\lambda_4(t) = 2\,r \sin h\omega(t_M - t)$
0	0	0	1	$\lambda_1(t) = 3\,\frac{\omega}{r} \left[1 - \cos h\omega(t_M - t)\right]$ $\lambda_2(t) = 0$ $\lambda_3(t) = -\frac{1}{r} \sin h\omega(t_M - t)$ $\lambda_4(t) = 3 - 2 \cos h\omega(t_M - t)$

Table 1. Solutions of the adjoint system for various terminal conditions.

The resulting variances of the round-off errors at an arbitrary time t_M are tabulated in Table 2. In a first approximation, the variances in the coordinates x_1, x_2 and x_4 increase proportional to the number of integration steps N, but the variance in x_2 increases proportional to N^3.

i	$\mathrm{Var}\left[u_i(t_M)\right]$	First Approximation $\omega t_M \ll 1$
1	$\sigma^2 \cdot \frac{4N}{t_M} \cdot \Big[\left(\frac{9}{4\omega} + \frac{1}{4\omega^3} + \frac{r^2}{\omega^3}\right) \mathrm{sin\,h}\,(2\omega t_M) - \left(\frac{12}{\omega} + \frac{8r^2}{\omega^3}\right) \mathrm{sin\,h}\,(\omega t_M) + \left(\frac{17}{2} - \frac{1}{2\omega^2} + \frac{6r^2}{\omega^2}\right) t_M\Big]$	$4N\sigma^2$
2	$\sigma^2 \cdot \frac{4N}{t_M} \cdot \Big[\left(\frac{9}{4\omega r^2} + \frac{1}{4\omega^3 r^2} + \frac{1}{\omega^3}\right) \mathrm{sin\,h}\,(2\omega t_M) + \left(\frac{18}{\omega r^2} - \frac{2}{\omega^3 r^2} + \frac{12}{\omega^3}\right) \mathrm{sin\,h}\,(\omega t_M) - \left(\frac{18 t_M}{r^2} + \frac{12 t_M}{\omega^2}\right) \mathrm{cos\,h}\,(\omega t_M) + \left(\frac{3\omega^2}{r^2} + 3\right) t_M^3 - \left(\frac{9}{2r^2} - \frac{3}{2} \cdot \frac{1}{\omega^2 r^2} + \frac{2}{\omega^2}\right) t_M\Big]$	$12\,N\,t_M^2\,\sigma^2$ or $12\,N^3 \Delta t^2\,\sigma^2$
3	$\sigma^2 \cdot \frac{4N}{t_M} \cdot \Big[\left(\frac{9\omega}{4} + \frac{1}{4\omega} + \frac{r^2}{\omega}\right) \mathrm{sin\,h}\,(2\omega t_M) + \left(\frac{1}{2} - 2r^2 - \frac{9\omega^2}{2}\right) t_M\Big]$	$4N\,\sigma^2$
4	$\sigma^2 \cdot \frac{4N}{t_M} \cdot \Big[\left(\frac{9\omega}{4r^2} + \frac{1}{4\omega r^2} + \frac{1}{\omega}\right) \mathrm{sin\,h}\,(2\omega t_M) - \left(\frac{18\omega}{r^2} + \frac{12}{\omega}\right) \mathrm{sin\,h}\,(\omega t_M) + \left(\frac{27\omega^2}{2r^2} - \frac{1}{2r^2} + 11\right) t_M\Big]$	$4N\,\sigma^2$

Table 2. The variances of the round-off errors after integrating the perturbation Equations (40) with Heun's method.

The error in the state transition matrix ${}_E\Phi(t_1, t)$ now can be characterized by

$$E\left[\delta_E \Phi\right] = 0$$

and

$$E\left[\delta_E \Phi^2\right] = \begin{bmatrix} \sigma_1^2 & 0 & \sigma_1^2 & \sigma_1^2 \\ \sigma_2^2 & 0 & \sigma_2^2 & \sigma_2^2 \\ \sigma_3^2 & 0 & \sigma_3^2 & \sigma_3^2 \\ \sigma_4^2 & 0 & \sigma_4^2 & \sigma_4^2 \end{bmatrix} \qquad (41)$$

where σ_i^2 is equal to $\mathrm{Var}\left[u_i(t_M)\right]$ and given in Table 2.

7.2 ERRONEOUS PREDICTION OF THE STATE VECTOR

Restricting ourselves again to an in-plane trajectory of a satellite, we have to integrate two second order differential equations of the form (30).

Setting

$$r = X_1 , \quad \theta = X_2 , \quad \dot{r} = X_3 , \quad \dot{\theta} = X_4 ,$$

we can write the equations of motion as a system of first order equations

$$\begin{aligned}
\dot{X}_1 &= X_3 \\
\dot{X}_2 &= X_4 \\
\dot{X}_3 &= - \frac{K}{X_1^2} + X_1 X_4^2 = f_1 (X_1, X_2, X_3, X_4) \\
\dot{X}_4 &= - \frac{X_3 X_4}{X_1} = f_2 (X_1, X_2, X_3, X_4)
\end{aligned} \tag{42}$$

and obtain the error equations similar to those in Section 7.1 :

$$\begin{aligned}
u_{1j} &= u_{1j-1} + \Delta t \, u_{3j-1} - 2 \epsilon_1 \\
u_{2j} &= u_{2j-1} + \Delta t \, u_{4j-1} - 2 \epsilon_2 \\
u_{3j} &= u_{3j-1} + \Delta t \left[\left(\frac{2K}{X_1^3} + X_4^2 \right) u_1 + (2 X_1 X_4) \, u_4 \right] - 2 \epsilon_3 \\
u_{4j} &= u_{4j-1} + \Delta t \left[\left(\frac{X_3 X_4}{X_1^2} \right) u_1 - \left(\frac{X_4}{X_1} \right) u_3 - \left(\frac{X_3}{X_1} \right) u_4 \right] - 2 \epsilon_4
\end{aligned} \tag{43}$$

The only difference is that the coefficients are functions of time here and not constants. This does not complicate matters too much because coefficient accuracy is not critical. Nominal values of the trajectory suffice. We may go even one step farther and restrict ourselves to an approximately circular orbit and determine the coefficients by setting

$$X_1 = r \cong \text{constant} \quad ,$$

$$X_3 \approx 0$$

$$X_4 = \omega^2 \cong \text{constant}$$

The equation system (43) reduces to (33) and can be solved with the elementary methods described in Section 7.1.

The error vector

$$\delta X(t_1) = \delta_M \Phi \delta X(t)$$

now can be characterized by

$$E\left[\delta X(t_1)\right] = 0$$

and

$$E\left[\delta X(t_1)\, \delta X^T(t_1)\right] = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix} \qquad (44)$$

where the σ_i^2 are equal to Var $\left[u_i(t_M)\right]$ and given in Table 2.

8. EXAMPLE: SATELLITE TRAJECTORY ESTIMATION BY THE ONBOARD DIGITAL COMPUTER

We assume an approximately circular orbit in the terrestrial field. The actual in-plane trajectory is perturbed around a precomputed reference trajectory ($r = 2.2 \cdot 10^7$ feet). Observations are made periodically and used in the spaceborne digital computer to estimate the current position and velocity in the trajectory plane according to Figure 2.

We are interested in finding the estimation error caused by the finite wordlength in the digital computer, and to separate the effect of

wordlength K,

observation interval t_M, and

number of integration steps, N, in one observation interval

on the error distribution. We know, that the error is mainly determined by the round-off errors in the numerical integration of the equations of motion. The variances of the accumulated error are given in general form in Table 2. Figure 3 shows the error propagation in the 4 coordinates $r, \theta, \dot{r}, \dot{\theta}$ for our example as a function of the integration time t_M and for a 0.5 seconds integration interval.

The error grows quite rapidly with N(or $t_M = N\Delta t$). The result is not in agreement with the commonly [6] assumed error growth proportional to N. It does not seem realistic to assume that the round-off errors grow proportional to N independently of equation structure, integration method, coordinate system, etc. It is noted in Table 2 that the error grows proportional to N (in agreement with ref. 6), but only for three of the four state variables and for a very short integration time.

The curves in Figure 3 are independent of the computation accuracy. If we multiply σ_i^{*2} with the variance, σ^2, of the round-off error in one integration step, we find the actual error variance σ_i^2 - Table 3.

We look now for a geometrical representation of our error covariance matrix, and introduce the ellipse of concentration. The ellipse (Figure 4) is defined as a curve enclosing all points which deviate from the reference point with a probability at most equal to P.

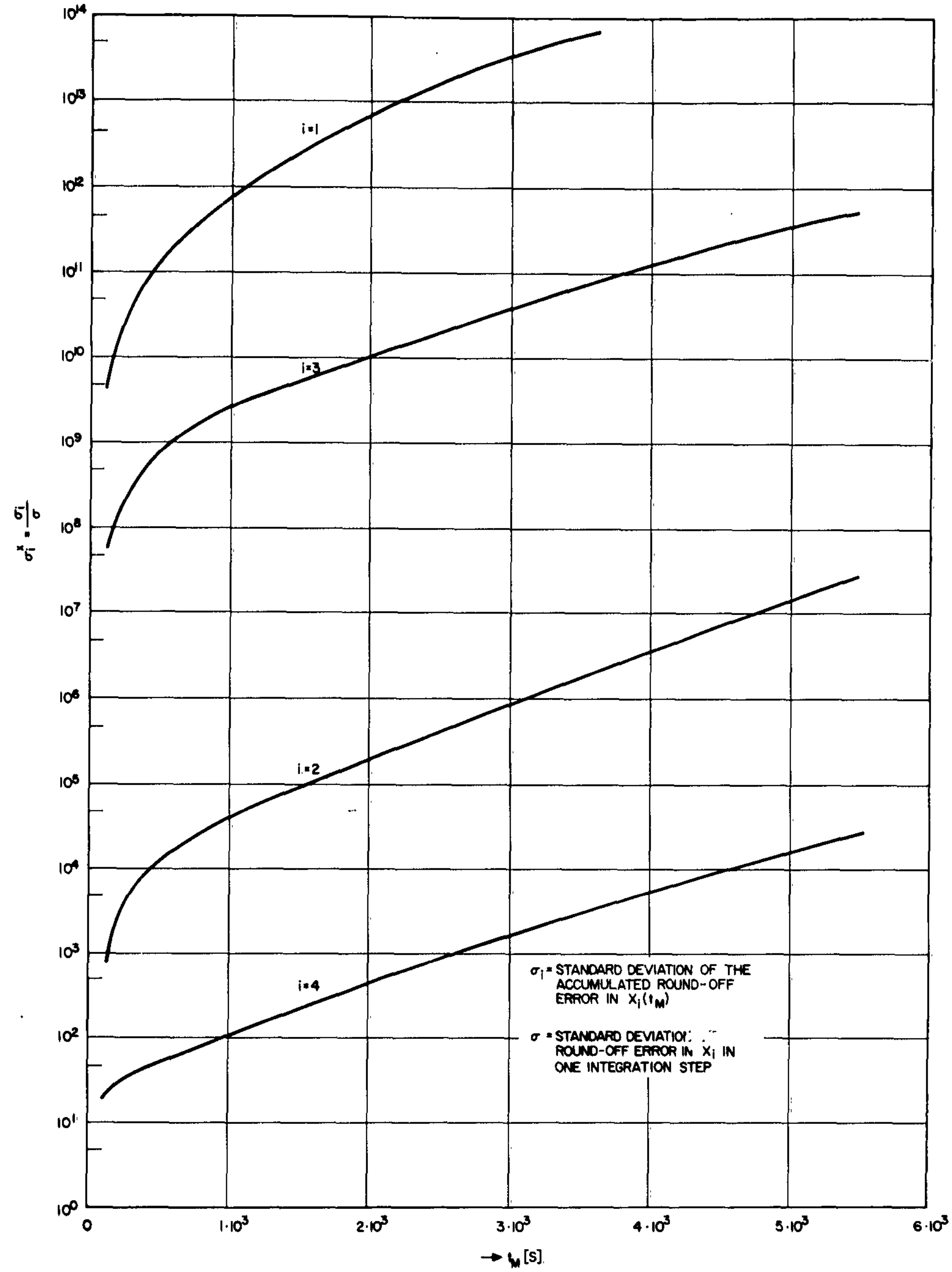

Figure 3. Accumulation of round-off errors during the numerical integration of a system of first order differential equations. Integration step 0. 5s.

Duration of Integration $t_M[s]$	Error in Coordinate	Word Length			Dimension
		K = 24	K = 28	K = 36	
600 (10 minutes)	X_1	2.495×10^{7}	$9.7 \times 10^{+4}$	1.47×10^{0}	$(\text{feet})^2$
	X_2	1.40×10^{-7}	0.545×10^{-9}	0.825×10^{-14}	$(\text{radian})^2$
	X_3	$4.81 \times 10^{+2}$	1.87×10^{0}	2.33×10^{-5}	$(\text{feet/s})^2$
	X_4	1.03×10^{-12}	0.4×10^{-14}	0.604×10^{-19}	$(\text{radian/sec})^2$
3600 (1 hr)	X_1	1.28×10^{12}	4.87×10^{9}	7.4×10^{4}	$(\text{feet})^2$
	X_2	1.55×10^{-3}	0.59×10^{-5}	7.93×10^{-11}	$(\text{radian})^2$
	X_3	1.96×10^{6}	7.44×10^{3}	1.133×10^{-1}	$(\text{feet/s})^2$
	X_4	3.38×10^{-9}	1.29×10^{-11}	1.96×10^{-16}	$(\text{radian/sec})^2$

Table 3. Variance of the accumulated round-off errors as a function of wordlength and integration time.

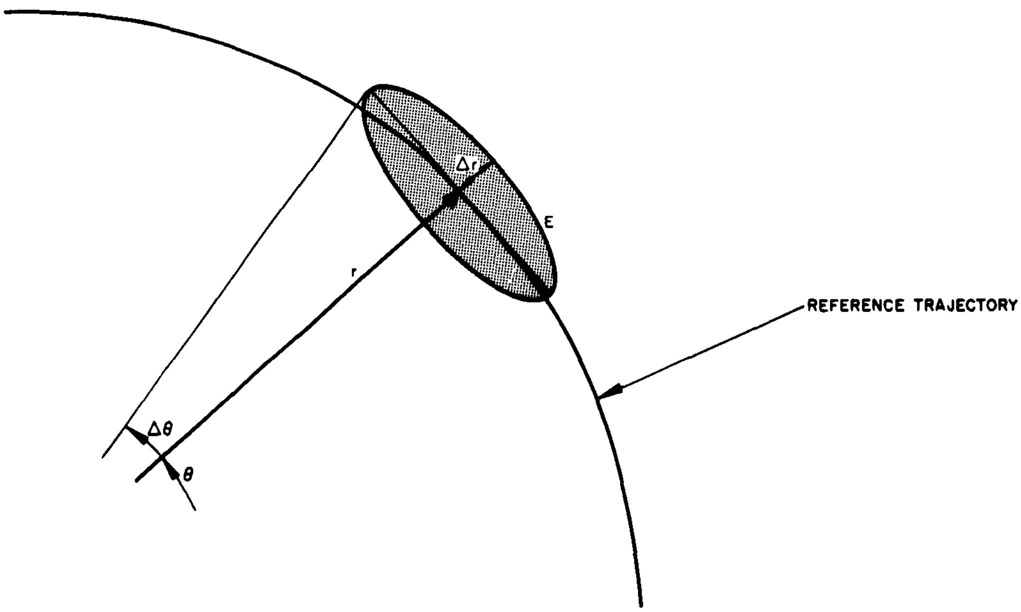

Figure 4. Ellipse of concentration as a performance measure for the digital computer in midcourse navigation.

Let x and y be cartesian coordinates with respect to the centre of the ellipse (on the reference trajectory)

$$x = \Delta r$$

$$y = r\Delta\theta$$

where Δr and $\Delta\theta$ are small,

then the equation of the homothetic ellipses is

$$\frac{x^2}{\sigma_1^2} + \frac{y^2}{r^2\sigma_2^2} = 2c^2 \tag{45}$$

where σ_1^2 and σ_2^2 are given in Table 2 and Table 3. The constant c^2 is χ^2 distributed with 2 degrees of freedom. The probability of falling outside of the ellipse is obviously given by

$$P = P(\chi^2 > 2c^2). \tag{46}$$

It is noted that in our example (Figure 4) the major axes of the ellipse always lie in the direction of r and perpendicular to r (downrange). But a coordinate transformation in equation system (42) would change length and direction of the axes.

The effect of wordlength on the error ellipse is shown in Table 4. The area covered by the ellipse grows in our particular example from $3.3 \cdot 10^2$ feet^2 to the discouraging magnitude of $1.7 \cdot 10^{10}$ feet^2 if we require a probability of 99.9 percent of not falling outside the ellipse. Obviously, we may generate any size ellipse and require any wordlength for the digital computer, if we can persuade ourselves to more or less optimistic probability figures (Figure 5).

K	a [feet]	b [feet]	F [feet^2]
24	1.8×10^4	3.05×10^5	1.7×10^{10}
28	1.12×10^3	1.91×10^4	2.1×10^7
30	0.282×10^3	0.47×10^4	1.35×10^6
32	0.7×10^2	1.19×10^3	8.4×10^4
34	1.76×10^1	2.98×10^2	5.3×10^3
36	4.4	7.45×10^1	3.3×10^2

Table 4. Half axis and area of the 99.9 percent error ellipse K = word length, observation interval 10 minutes, approximately circular orbit with 2.2×10^7 feet radius.

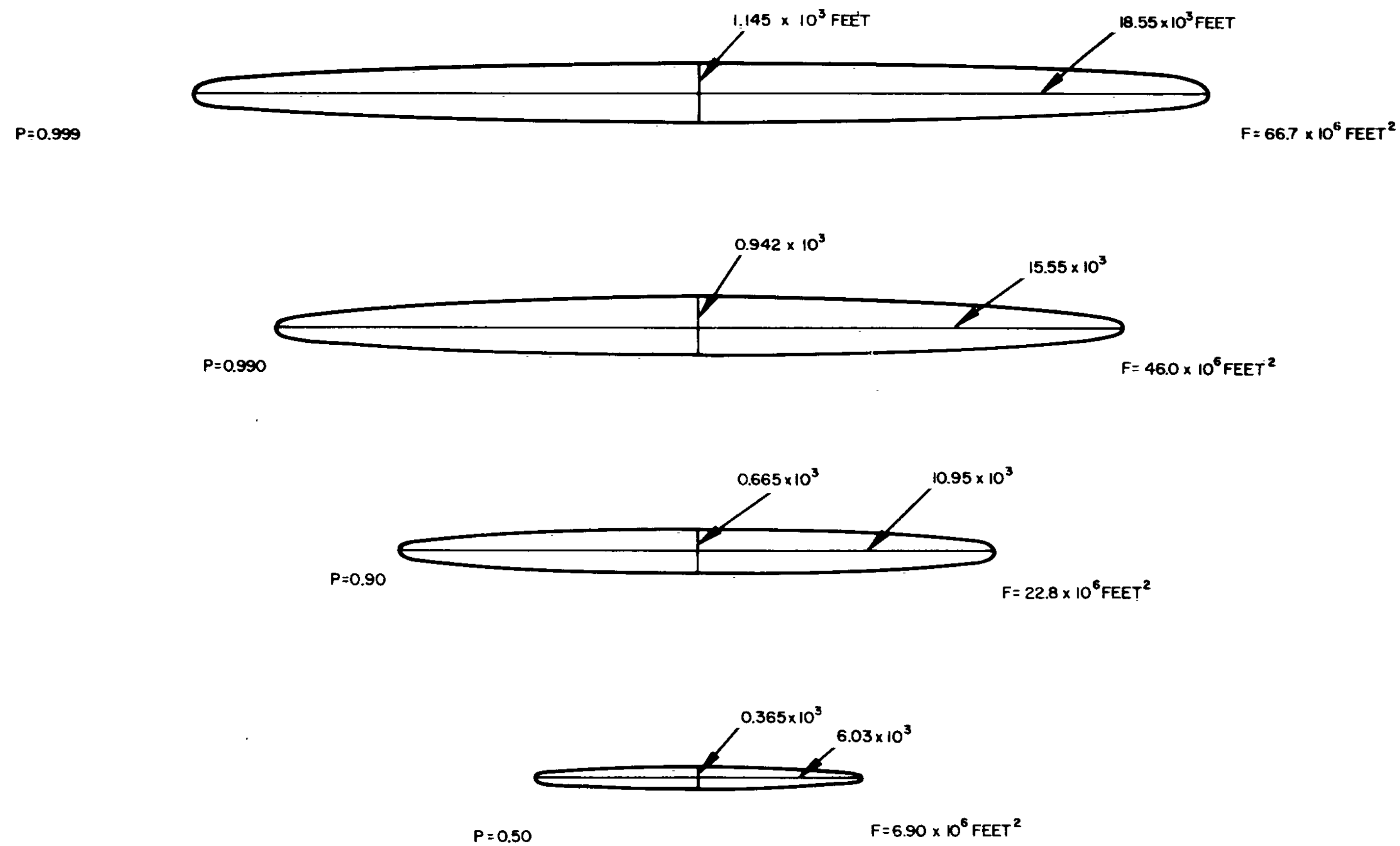

Figure 5. Error ellipses for varies error probabilities. Wordlength 28 digits, observation interval 10 minutes, approximately circular orbit with 2.2 x 10^7 feet radius.

9. SUMMARY

The error ellipse (ellipse of concentration) has been introduced as a performance measure, and its axes and area have been expressed in terms of the computer wordlength, the time interval between observations, and the number of integration steps for a particular class of estimation systems. Characteristic of the system is that the equations of motion are solved by numerical integration and not analytically. The conclusions about the growth of the estimation error are, therefore, rather pessimistic. A computer designed to operate satisfactorily in our example certainly would also be appropriate for more elegant navigation systems.

As a sideline to our main problem, it was necessary to obtain some information about the round-off error propagation in numerical integration of first order differential equations (second order equations can readily be reduced to first order equations). The error moments have been found by means of a first order perturbation method and have been expressed in terms of the adjoint solutions of the error equations. In general, the error variances grow with some power of N. The power depends on the integration time, number of variables in the equation system, and on the form of the equations (coordinate system). The coefficients of the differential equations have been made approximately constant. Otherwise, the adjoint equations cannot be integrated in closed form.

If we are satisfied with first approximations of the estimation error and if we may assume convergence of the regular estimation error to a moderately small equilibrium (steady state), then the problem of finding a performance measure for the digital computer during midcourse navigation is reduced to the task of determining the round-off error propagation during numerical integration of the equations of motion and perturbation. Table 2 gives closed form expressions for the accumulated round-off error variances assuming approximately circular orbit and in-plane two body motion (Section 7.1 and 7.2).

The effect of limited computer accuracy on the estimation error can be determined with the Formulas (28) and (29). Both formulas are approximations. Neglected are:

a. Propagation of round-off errors from previous observation intervals against the error contributions of the last interval. Round-off errors from previous observation intervals would contribute only in pathological cases where $X(t_i)$ decreases so rapidly with t_i that $2\delta\phi^2 X(t-2)$ is not neglegible against $\delta\phi X(t-1)$.

b. The effect of the erroneous filtering operation $\Delta^*(t) + \delta\Delta^*$. It can safely be done only, if the regular estimation error $\tilde{x}(t)$ converges with the increasing number of measurements: $\lim_{N\to\infty} E\left[\tilde{x}\ \tilde{x}^T\right] = 0$

The errors in the state transition matrices ${}_M\Phi$ or ${}_E\Phi$ are associated in (28) and (29) with the perturbation vector $x(t)$ or with the state vector $X(t)$. Even very small $\delta\phi$ may contribute significantly to the estimation error, if $x(t)$ or $X(t)$ are sufficiently large. A typical example is the transition matrix of injection errors for a circular orbit[12]. The matrix contains elements which grow with t. Unbounded perturbation components $x_i(t)$ and unbounded solutions of the variance equation must be expected.

We notice, that computation noise may increase the uncertainty in our knowledge of the system state, even if a large number of observations have been made. We found that the covariance matrix of the computation noise appears to be additive to the covariance matrix of the random force $u(t)$ in (18). We may define, therefore, a generalized covariance matrix

$$Q^*(t) = Q(t) + \left[P_D(t) - P(t)\right]$$

and recognize that $Q^*(t)$ may be unbound. We conclude that theorem 4 in reference 2 does not apply for dynamic systems which are corrupted by computation noise.

10. REFERENCES

1. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems", Journal of Basic Engineering, ASME, March 1960

2. Kalman, R. E., Bucy, R. S., "New Results in Linear Filtering and Prediction Theory", Journal of Basic Engineering, ASME, March 1961

3. Von Neumann, John, Goldstine, H. H., "Numerical Inverting of Matrices of High Order", Bulletin of the American Mathematical Society, Vol 53, Nr 11, November 1947

4. Rademacher, Hans A., "On the Accumulation of Errors in Processes of Integration on High-Speed Calculating Machines", Proc. Symposium on Large Scale Digital Computing. Ann. Computing Lab. Harvard 16 (1948)

5. Smith, Gerald L., Schmide, Stanley F., McGee, Leonard A., "Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity On Board A Circumlunar Vehicle", NASA Tr R 135, 1962

6. Conte, Samuel D., "The Computation of Satellite Orbit Trajectories", Advances in Computers, Vol 3, 1962, Academic Press

7. Henrici, P., "Discrete Variable Methods in Ordinary Differential Equations", Wiley, New York, 1961

8. Holick, A., "Analysis of Noncatastrophic Failures in Digital Guidance Systems", IRE Transaction on Electronic Computers, August 1963

9. Holick, A., "Word Split Technique as a Means to Increase the Failure Tolerance Level in Core Memories", HAC Report No. SRS-622, February 1963

10. Holick, A., "Effect of Degraded Operational Modes on the Accumulation of Errors in Numerical Integration", HAC Report No. SRS-596, December 1962

11. Terris, I., "Techniques for Improving Computer Reliability by Logical Design (Self-Reorganizing Machines)", HAC Report, IDC 2924.01/10, May 10, 1963

12. Wisneski, M. L., "Error Matrix for a Flight on a Circular Orbit", ARS Journal, September 1962

APPENDIX A

GENERAL ADJOINT DIFFERENTIAL EQUATIONS FOR THE ROUND-OFF ERROR

The equations of motion are generally given by a nonlinear system of second order differential equations. We write it as a system of first order differential equations:

$$
\begin{aligned}
\dot{X}_1 &= X_4 \\
\dot{X}_2 &= X_5 \\
\dot{X}_3 &= X_6 \\
\dot{X}_4 &= f_1 (X_1 , X_2 , \ldots , X_6) \\
\dot{X}_5 &= f_2 (X_1 , X_2 , \ldots , X_6) \\
\dot{X}_6 &= f_3 (X_1 , X_2 , \ldots , X_6)
\end{aligned}
\tag{47}
$$

and linearize by Taylor series expansion around the reference trajectory. Introducing the perturbation coordinates

$$x_i = X_i - X_{i,R} , \quad i = 1, \ldots , 6$$

we obtain for the perturbation equations:

$$
\begin{aligned}
\dot{x}_1 &= x_4 \\
\dot{x}_2 &= x_5 \\
\dot{x}_3 &= x_6 \\
\dot{x}_4 &= \sum_{i=1}^{6} \frac{\partial f_1}{\partial X_i} x_i = \sum_{i=1}^{6} F_{1i} (x_i)
\end{aligned}
\tag{48}
$$

$$\dot{x}_5 = \sum_{i=1}^{6} \frac{\partial f_2}{\partial X_i} x_i = \sum_{i=1}^{6} F_{2i}(x_i)$$

$$\dot{x}_6 = \sum_{i=1}^{6} \frac{\partial f_3}{\partial X_i} x_i = \sum_{i=1}^{6} F_{3i}(x_i)$$

Numerical integration on a digital computer using Heun's method leads to round-off errors which can be characterized in step j by the error equations:

$$
\begin{aligned}
u_{1j} &= u_{1j-1} + \Delta t\, u_{4j-1} - 2\epsilon_1 \\
u_{2j} &= u_{2j-1} + \Delta t\, u_{5j-1} - 2\epsilon_2 \\
u_{3j} &= u_{3j-1} + \Delta t\, u_{6j-1} - 2\epsilon_3 \\
u_{4j} &= u_{4j-1} + \Delta t \sum_{i=1}^{6} \frac{\partial F_{1i}}{\partial x_i} u_{ij-1} - 2\epsilon_4 \\
u_{5j} &= u_{5j-1} + \Delta t \sum_{i=1}^{6} \frac{\partial F_{2i}}{\partial x_i} u_{ij-1} - 2\epsilon_5 \\
u_{6j} &= u_{6j-1} + \Delta t \sum_{i=1}^{6} \frac{\partial F_{3i}}{\partial x_i} u_{ij-1} - 2\epsilon_6
\end{aligned}
\tag{49}
$$

where the ϵ_i designate the error in one integration step and $u_{ij} = x_i(t_j) - \overline{x}_i(t_j)$.

We solve the Equation system (49) by means of the adjoint equations

$$\lambda_{1j} = \lambda_{1j-1} - \Delta t \sum_{K=1}^{3} \frac{\partial F_{K1}}{\partial x_1} \lambda_{(K+3),\, j}$$

$$\lambda_{2j} = \lambda_{2j-1} - \Delta t \sum_{K=1}^{3} \frac{\partial F_{K2}}{\partial x_2} \lambda_{(K+3),\, j}$$

$$\lambda_{3j} = \lambda_{3j-1} - \Delta t \sum_{K=1}^{3} \frac{\partial F_{K3}}{\partial x_3} \lambda_{(K+3),\, j}$$

$$\lambda_{4j} = \lambda_{4j-1} - \Delta t \left\{ \lambda_{1j} + \sum_{K=1}^{3} \frac{\partial F_{K4}}{\partial x_4} \lambda_{(K+3)\, j} \right\} \tag{50}$$

$$\lambda_{5j} = \lambda_{5j-1} - \Delta t \left\{ \lambda_{2j} + \sum_{K=1}^{3} \frac{\partial F_{K5}}{\partial x_5} \lambda_{(K+3)\, j} \right\}$$

$$\lambda_{6j} = \lambda_{6j-1} - \Delta t \left\{ \lambda_{3j} + \sum_{K=1}^{3} \frac{\partial F_{K6}}{\partial x_6} \lambda_{(K+3)\, j} \right\}$$

and obtain the variance of the round-off error in the coordinates $x_i(t_M)$ at time t_M readily with the argumentations in Section 7. 1:

$$\mathrm{Var}\left[\sum_{i=1}^{6} \lambda_i(t_M)\, \mu_i(t_M) \right] = \frac{4\sigma^2}{\Delta t} \cdot \int_0^{t_M} \sum_{i=1}^{6} \lambda_i^2(t)\, dt. \tag{51}$$